Improving Data Discovery and Access through Interoperable System in Climate.gov

Sudhir Raj Shrestha (sudhir.shrestha@noaa.gov)
Jebb Stewart (jebb.q.stewart@noaa.gov)
Jason Marshall (jason.marshall@noaa.gov)

What is "Interoperability for Climate.gov"?

For the Climate.gov "Interoperability is an interface that allows users to find, display, manipulate, and (where applicable) download NOAA's and its partners' climate data products that are stored in and served from different data centers."

Why do we want it?

* Expansion of data use and efficiency and reach out to larger group of the people

* Increase/improve accessibility of NOAA data and cross the bridge of data formats

* One stop source of Climate Information

* Encourage Standards that increases the larger pools of the users' with same investment in data

* Standards will limit the diversity of data providing scheme and will create less complexity in the data management with improved implementation.

Impacts

Framework built towards Standards, NOT Data.

Important Because:

- Data is ready for action. Services model facilitates agile capabilities. Services can be combined or reused quickly.

- Any data available through framework can be operated on or combined with other data. Integrated standardized formats and access.

- New and existing systems have access to wide variety of data. Any new data added, easily incorporated with minimal to no changes required.

Interoperability Requirements

Format Agnostic.

Platform Agnostic

Owner/Physical Location Agnostic.

Preview Capabilities.

Semantics/Ontology/Vocabulary.

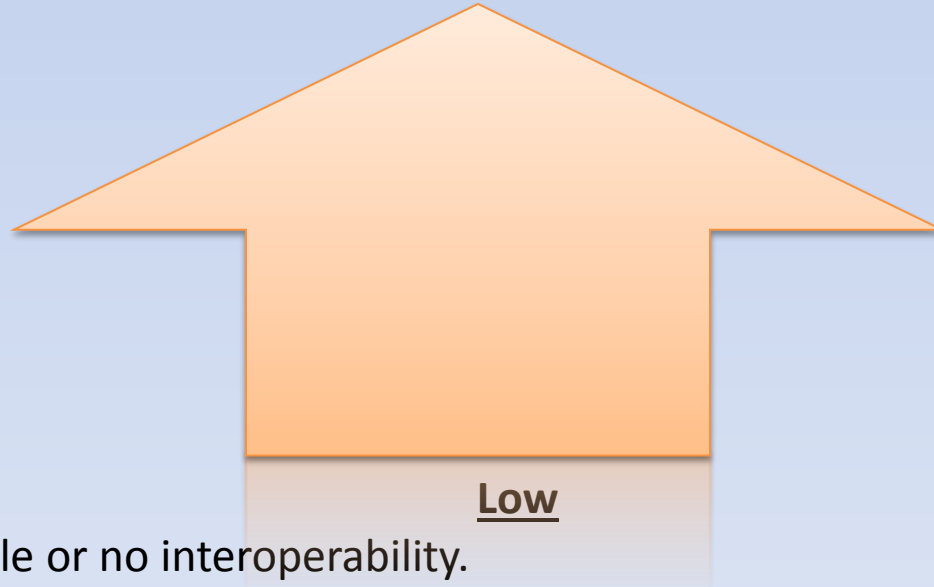Machine to Machine Communication.

Complete Metadata

Measuring Interoperability

High

Extensive interoperability.
Little human interpretation and intervention required.
Simple configuration rather than custom coding.

Low

Little or no interoperability.
Significant human interpretation and intervention required.
Extensive custom coding.

Climate.gov
science and services for society

Interoperability Readiness Levels

- Measureable indicators...

	Capability Enablement	Discovery	Access	Understanding	Data	Standards
Level 1	Data from different sources cannot be used together	Probably none, hard coded or inaccessible catalog interface	Not modular components (part of a larger application), platform specific, undocumented, no distributed access, closed/restricted source (not open source)	Content of data is not directly comparable to any other data	Data in unknown or undocumented formats with little or no auxiliary content information available	Individual standards.
Level 2	Human use of data from different sources using different code for each	REST-style access to form interface (via scraping)	Proprietary and complicated dependencies, strict platform dependencies, limited documentation, no discovery (registry)	Some parts of data may be comparable to other data only through informal human to human interaction	Data in documented formats with little or no auxiliary content information available	
Level 8	Human-triggered incorporation of novel data and services into applications	Services discoverable in global registries of services with complete syntactic information	Discipline/Domain-specific ontology support using recognized semantic tools	Semantic agreement on content based upon community-accepted ontologies	Standard data types in syntactically self-describing formats, quality, applicability, etc. information partly semantically captured	
Level 9	Automatic discovery and incorporation of novel data and services into applications with no human intervention	Services discoverable in global registries with complete syntactic and semantic information	AI capability. Completely automated mediation of services.	Semantic agreement on content based upon universally accepted ontologies	Data, its quality, realms of applicability, etc. fully self-described both syntactically and semantically	International standards.

Climate.gov
science and services for society

Interoperable System Architecture

CORS-= Cross Origin Resource Sharing. Open to all domain
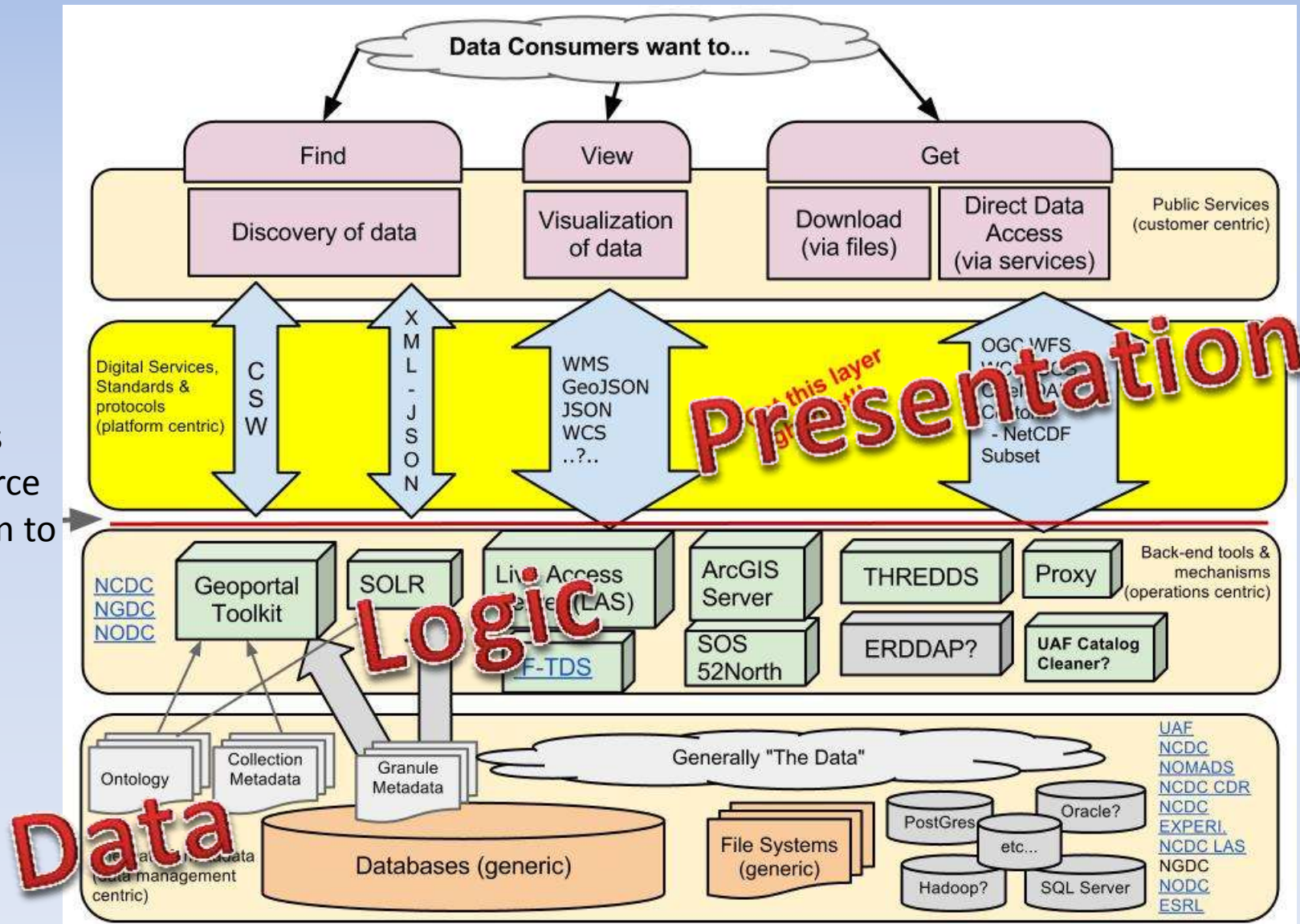
Prototyping for select Use Case

- Built a proof of concept Data Interoperable Platform

- The built system is "file format agnostic," meaning the pilot system will locate and display the data regardless of what format they're archived

- The web based client was developed using javascript libraries from OpenLayers and JQuery. OpenLayers library provides javascript utilities to interact with a variety of data and metadata services. JQuery provides utilities to construct the layout of the web page itself

- The codes are available to download and fork out in github: https://github.com/ClimateData/interoperability

Demo of Climate.gov Interoperability Prototype

http://www.esrl.noaa.gov/neis/service/cpodi/index.html

* This link only works after disabling browser cross domain security

Challenges

- Mixed environments
- Multiple data types and Data Discovery, Access
- GIS Services (ESRI + Open Source)
- Technological change
- Diverse standards adoption

Findings

Metadata incomplete, difficult to maintain.
- Create tools to improve and automate metadata creation.

Data not easily discoverable. Often requires prior knowledge to answer basic questions: who, what, where, how?
- Metadata, Metadata, Metadata.

Limited details of what this data should look like?
- Preview or other information of how should data be visualized.
- Open Geospatial Consortium (OGC) Styled Layer Descriptor (SLD) and Symbology Encoding (SE).

Incomplete or lack of adherence to standards.
- Analyze and provide feedback to data providers.
- Do services meet standards?

Cross Domain Content is restricted
- Implement CORS (Cross-Origin Resource Sharing)

Next Steps…

- Continue towards Interoperability

- Standards and policy

- Compliant Metadata

- Data served through GIS Services

- Advocacy & outreach

- Solicit suggestions on our proposed Interoperability System Architecture.

Acknowledgement

- Steve Ansari (NOAA-NCDC)

- Kevin O'Brien (NOAA-PMEL)

- David Herring (NOAA-CPO)

- Mark Phillips (UNCA)

- Micah Wengren (NOAA)

- Mike Halpert (NOAA-CPC)

Any Questions ?

Climate.gov Data Team

sudhir.shrestha@noaa.gov

Jebb.q.stewart@noaa.gov